Integrated Cortical-Hippocampal Neural Architecture of Episodic Memory

New computational model improves fidelity to neurobiology

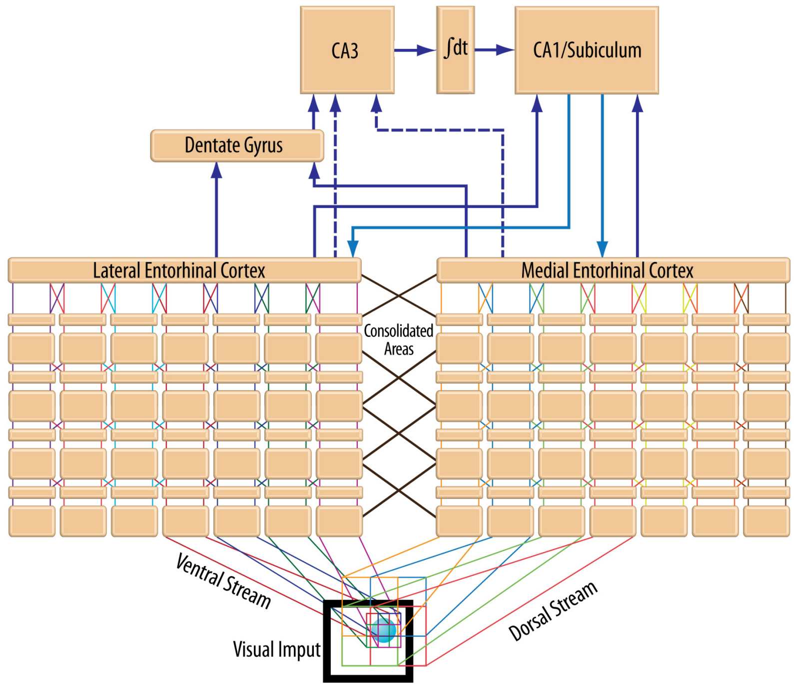

Figure 1: Simplified Sandia model.

For more information:

Technical Contact:
Michael Bernard
505-845-0815
mlberna@sandia.gov

Science Matters Contact:
Alan Burns
505-844-9642
aburns@sandia.gov

In cooperation with Boston University, the University of Illinois at Urbana-Champaign, and the University of New Mexico, Sandia is developing a neurologically plausible, artificial neural network architecture of episodic memory and recall, modeled after cortical-hippocampal structure and function.

The neural network architecture is based on research concerning multiple memory systems in the brain, and makes a distinction between declarative memory (for facts and events), and procedural or non-declarative memory (supporting the acquisition and expression of skills). Declarative memory is fundamentally relational, involving representations of the relations among the constituent elements of experience. This includes both episodic memory, constituting the binding together in memory of the who, what, where, and when of events, and semantic memory, involving binding in memory of information constituting the structured knowledge we acquire about the world.

Progressing from previous architecture versions in which the entire hippocampus was represented by a conjunctive grid, Sandia has instead developed individual modules to represent the primary regions of the hippocampus. This architecture is depicted in Figure 1. The main distinction between this approach and others is that it focuses more on the manipulation of memory relations rather than the encoding of specific events or neuronal activity. This is biologically plausible since it bases the architecture on estimates of the relative number of neurons and neuronal connections, but does not attempt to model the interaction between neurons individually. Building upon accepted

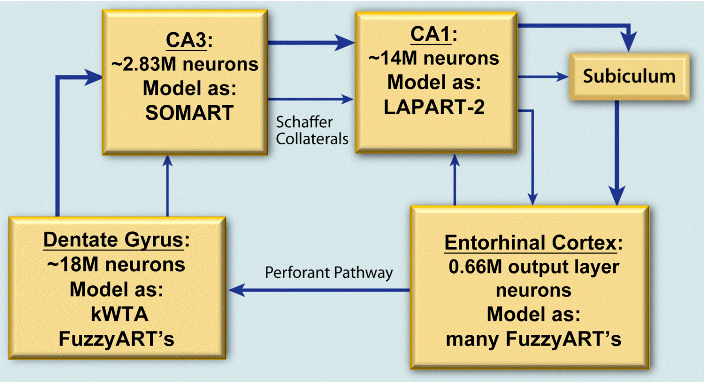

Figure 2: The relative number of human hippocampal neurons and the Sandia hippocampal loop representation.

hippocampus sub-region functionality, the researchers have implemented a computational architecture with conceptual processing based upon variants of Adaptive Resonance Theory (ART) for artificial neural networks. The pre-medial temporal lobe sensory cortex, and the entorhinal cortex (EC) are represented by layers of modified fuzzy-ART modules. The fuzzy-ART capabilities are extended by equipping these modules with the capability to encode temporal semantic data. Individually, these temporally integrated adaptive resonance theory (TIART) modules are capable of encoding categorical representations of their given input vectors over time. By combining layers of TIART modules, categories of categories may be formed to represent larger semantic concepts. In neuroanatomy, the EC receives multimodal sensory inputs. The model here simulates visual processing in which a dorsal stream containing contextual information and a ventral stream of focal information convene at the EC before entering the hippocampus.

Within the hippocampal representation in the model, each of the primary regions is represented by a different ART variant selected to achieve the particular functionality of the individual region. The relative size of each module is scaled in accordance with approximate human neuroanatomy. The particular module implementations as well as their respective size scaling may be seen in Figure 2.

The dentate gyrus (DG) receives the conjoined multimodal sensory signals from the EC. It performs pattern separation on this abundance of sensory information to produce sparse output activation, which ensures different semantic concepts are given unique encoding. This sparse output pattern from the DG serves as input to CA3. Functionally, CA3 assists with episodic binding through auto-association. In Sandia's model, this module is capable of mapping semantically similar inputs to proximate topological regions. In effect, related concepts are clustered together to help associate episodic memories. CA1 is involved in forming sequences of relations and connecting these episodic encodings back to the original sensory inputs from the EC. This ability to link sequences allows for temporal packaging of episodes. To handle the sequencing of associations, the researchers temporally integrate the inputs passed into the model's conjoined representation of CA1 and subiculum. A laterally primed adaptive resonance theory (LAPART) module represents the conjoined CA1 and subiculum regions. The LAPART module learns to associate CA3 encodings with the original unaltered EC inputs. This allows the LAPART module to complete the hippocampal loop and propagate temporal sequences back to EC and eventually to cortex for long term storage.

The progression of the computational model is driven by attempts to improve fidelity in relation to neurobiology. The approach has been to model neuroanatomy and in doing so, Sandia has demonstrated the ability to model elements of cognitive behavior such as familiarity and recognition. As a result of improving the hippocampal model the researchers are also able to create automatic associations of various semantic concepts (see Figure 3). In general, the artificial neural network computation model presented here processes sensory inputs and in effect is capable of exhibiting qualitative memory phenomena such as auto-association of episodic memory concepts.

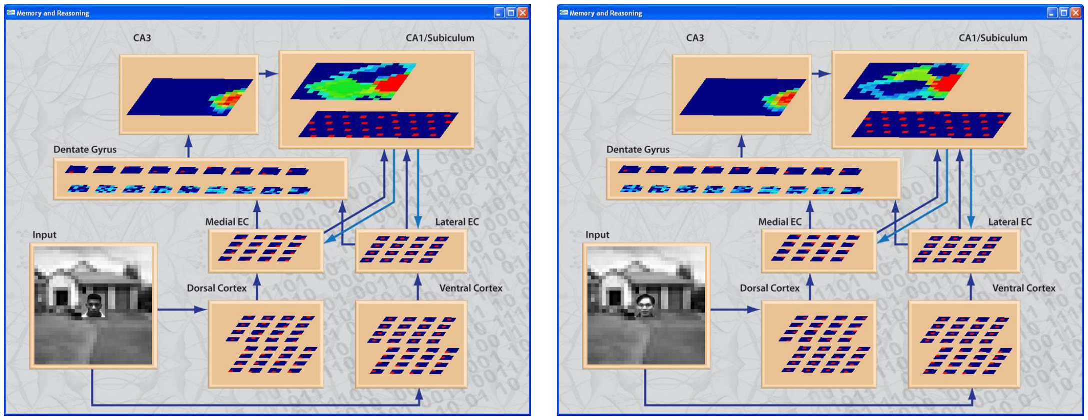

Figure 3: The current architecture can auto-associate and differentiate between different individuals with the same background.

Sandia National Laboratories

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000. SAND2009-5577P

National Nuclear Security Administration

09/2009